

Sieben Schritte für die Datenqualität

Kunden sind die Grundlage jedes Unternehmens. **Genauere Kontaktinformationen der Kunden steigern nicht nur die effiziente Kommunikation, sondern sparen auch Kosten.** So finden Sie eine geeignete Lösung zur Sicherstellung der Datenqualität.

Text _ Bud Walker

Stellen Sie sich ein Szenario vor, in dem mehrere Call-Center-Mitarbeiter Adressinformationen von Anrufern manuell in eine Kundendatenbank einpflegen. Wussten Sie, dass bis zu 20 Prozent dieser Kontaktdaten mangelhaft sind?

Analysten haben die langfristigen Kosten einer solchen ungenauen Datenerfassung untersucht. Ihre Schlussfolgerungen über die tatsächlichen Kosten von falschen Daten sind sehr aufschlussreich. Sie untersuchten systematisch die Kosten vom Point of Entry bis hin zum Versand der Produkte, die Rechnungsstellung und die Konsolidierung von Kundendaten für Data-Mining-Projekte. Es wurde festgestellt, dass es im Durchschnitt etwa einen Dollar kostet, um eine Kontaktinformation zu überprüfen, bevor diese auf die Kundendatenbank übermittelt wird. Das dient zum Schutz und unterbindet die Speicherung von

ungeprüften Adresseinträgen. Dies sorgt für saubere Adressen von Anfang an. Diese Kalkulation enthält die Kosten einer Lösung zur Adressvalidierung, als auch die Kosten von Mitarbeiterstunden sowie die Kosten für den Betrieb der IT-Infrastruktur pro Datensatz.

Karteileichen aussortieren

Wurde die Adresse abgespeichert, um diese in der Stapelverarbeitung (Batch) überprüfen zu lassen, belaufen sich die Kosten im Schnitt auf rund zehn Dollar je Datensatz. Die Erklärung hierfür ist, dass einige Adressen so fehlerhaft eingetragen wurden, dass keine noch so intelligente Software die Adresse im Nachgang korrigieren kann. Dies sind weder Sales Leads noch ertragreiche Kundenbeziehungen – sondern lediglich Karteileichen. Darüberhinaus berücksichtigten die Experten hierbei

auch die zusätzlichen Kosten der Software-Rechenzeit, den Aufwand zur Aufbereitung und Vereinheitlichung der Daten aus unterschiedlichen Formaten und die Einstellungszeit, um Batchläufe individuell zu verarbeiten. Schließlich wurde aufgezeigt, dass die Kosten des Nichtstuns bei rund 100 Dollar pro Datensatz liegen. Die Gründe liegen auf der Hand: Verspätete Sendungen, Retouren, Fakes bzw. Betrugsversuche, ineffiziente Marketingkampagnen und entgangene Umsatzchancen.

Datenqualität entsteht nicht durch Zufall. Sie erfordert einen Plan, der alle aktuellen Schwachstellen der Daten im Unternehmen aufdeckt, die gewünschten Ergebnisse liefert und mit den vorhandenen Ressourcen erreicht werden kann. Darüberhinaus sollte dieser Plan auch künftige Bedürfnisse berücksichtigen. Eine Datenqualitätslösung sollte die Möglichkeit bieten, mit dem Unternehmen und seinen Anforderungen zu wachsen.

Zunächst sollte das Unternehmen erkennen und sich darüber bewusst werden, dass ein Problem mit der Qualität von Daten besteht. Dies erfordert eine strenge Prüfung der Kosten, die durch ungenaue Daten entstehen. Es sollten Kosten einbezogen werden, die direkt in Verbindung mit schlechten Adressdaten stehen (zum Beispiel Versandkosten, Retouren bzw. Postrückläufer), aber auch entgangener Umsatz und Mitarbeiterstunden, die zur Problembehandlung aufgebracht wurden – was durch eine Datenqualitätslösung hätte verhindert werden können. Dann sollte die Bestimmung eines Data-Quality-Verantwortlichen erfolgen (oder

→ INFO KRITERIEN DER DATENQUALITÄT

Der Begriff Datenqualität kann für unterschiedliche Arten von Daten stehen. Bei Kundendaten trägt jeder Datensatz, der die genannten Qualitätsstandards nicht erfüllt, zu Risiken und unnötigen Kosten bei.

→ **Die Daten sind richtig:** Straße, Postleitzahl und Land stimmen überein.

→ **Die Daten sind aktuell:** Regelmäßiger Abgleich gegenüber Referenzdaten.

→ **Die Daten sind vollständig:** Straße, Postleitzahl und Land sind enthalten.

→ **Die Daten sind nicht redundant:** Dubletten werden identifiziert und gegebenenfalls zusammengeführt.

→ **Die Daten sind standardisiert:** Die Schreibweise ist einheitlich (zum Beispiel einheitlich Strasse, Straße oder Str. oder einheitlich Nummer, Nr oder NR).

eines Data-Quality-Teams), welcher für die Ausführung der Datenqualitätsinitiative zuständig sein wird. Diese Person ist für die Identifizierung und die möglichen Lösungsansätze des Problems zuständig. Nachdem man sich für einen Ansatz entschieden hat, sollte diese zentrale Kontaktperson für Datenqualität die Umsetzung der entsprechenden Initiativen durchführen.

Qualitätsprobleme bewerten

Der erste Punkt des Verantwortlichen sollte es sein, die eigentliche Ursache der Datenqualitätsprobleme im Unternehmen zu identifizieren. Es sollte herausgefunden werden, wo Qualitätsprobleme entstehen und wie diese den Weg ins Unternehmen finden.

»DATENQUALITÄT ERFORDERT EINEN PLAN, DER ALLE AKTUELLEN SCHWACHSTELLEN DER DATEN IM UNTERNEHMEN AUFDECKT.«

Die Entscheidung, sich des Themas Datenqualität anzunehmen, erfordert außerdem eine sorgfältige und realistische Einschätzung der eigenen technologischen Möglichkeiten und der Ressourcen, die innerhalb des Unternehmens zur Verfügung stehen. Folgende Fragen sind zu klären: Falls schlechte Daten online über die Website generiert werden, ist es möglich, Technologien zu integrieren oder wird man die Website umrüsten müssen? Wie groß ist das Adressvolumen? Benötigt man eine lokale Lösung für große Datenmengen oder eine kostengünstigere Web-Service-Adressvalidierung für kleine bis mittelgroße Datenvolumina? Hat das Unternehmen Entwickler zur Integration von Datenqualitätstools in eine bestehende Anwendung? Ermöglicht die vom Unternehmen verwendete Software die Einbindung von externen Schnittstellen oder ist es ein geschlossenes System? Wenn es ein geschlossenes System ist: Kann eine Standalone-Software dem Bedarf des Unternehmens gerecht werden? Was wäre der beste Ansatz um sicherzustellen, dass alle Adressdaten

einen zentralen Datenqualitätsprozess durchlaufen, bevor diese in irgendeiner Weise verwendet werden?

Sobald das Ausmaß des Problems und die verfügbaren Ressourcen erkannt wurden, sollte die verantwortliche Person für Data-Quality verschiedene am Markt erhältliche Datenqualitätslösungen in Betracht ziehen, deren Produkte analysieren und beurteilen. Die Evaluierung sollte einen Download und die beschränkte Nutzung einer kostenfreien Testversion der Lösung umfassen. Außerdem sollte die Möglichkeit geboten sein, mit Referenzkunden des Softwareanbieters für einen Erfahrungsaustausch in Kontakt zu treten. Nachdem der Vergleich mit möglichst mehreren Lösungen abgeschlossen ist, gibt der Data-Quality-

Experte seine Empfehlungen an die Geschäftsleitung hinsichtlich der besten bzw. geeignetsten Lösung für das Unternehmen und das angedachte Einsatzszenario ab.

In dieser Phase wird die empfohlene Datenqualitätslösung in der Praxis in einem begrenzten Umfang für eine Probelaufzeit eingebunden. Der Data-Quality-Mitarbeiter untersucht die Ergebnisse unter Berücksichtigung der Kosten und des Arbeitsaufwandes – und vergleicht diese mit einem ähnlichen Zeitraum vor dem Einsatz der Datenqualitätslösung.

Nach der Probephase werden die Ergebnisse aufbereitet und an die Geschäftsleitung zurückgemeldet. An diesem Punkt kann die Wirksamkeit der Datenqualitätsinitiative ausgewertet und – falls notwendig – können etwaige Anpassungen vorgenommen werden, bevor die Lösung im vollen Umfang innerhalb der gesamten Organisation eingebunden wird.

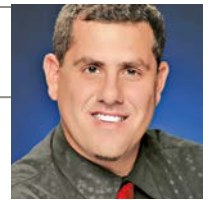
Es erfordert keine großen Investitionen für eine Lösung, um direkt während der Datenerfassung zu prüfen und zu ge-

AUTOR

Bud Walker

ist Director of Data Quality Solutions bei Melissa Data. Er ist für den Software Lifecycle und die Beratung zu Data-Quality-Lösungen und Services verantwortlich.

→ www.melissadata.de



währleisten, dass ausschließlich valide Kundenkontakte vorliegen.

Ein weiterer Vorteil einer Adressprüfung in Echtzeit ist es, Kundendaten zu verbessern oder mit Informationen anzureichern. Durch das Anhängen von wertvollen Informationen, wie zum Beispiel Geokoordinaten, erhalten Firmen standortbasierte Informationen zu Kunden und Interessenten. Durch das Anfügen von sozio-demografischen Merkmalen in Form von Lifestyle-Daten, können Kunden segmentiert werden, um gezieltere Kampagnen durchzuführen – und in einigen Ländern profitiert man bei geprüften Adressen von Postrabatten. Unternehmen erzielen Verbesserungen im Kundenservice und stärken die Kundenbindung. Der Erfolg im Direktmarketing zeigt sich an höheren Response-Raten, die schneller und effizienter erreicht werden.

Unter dem Strich gilt: Es ist wesentlich ökonomischer, eine Lösung zur Datenvalidierung am Anfang der Datensammlungskette zu implementieren.

redaktion@acquisa.de

•]

SUMMARY

- **ADRESSPRÜFUNG** Vorteil einer Adressprüfung in Echtzeit ist es, Kundendaten zu verbessern.
- **DATENVALIDIERUNG** Es ist ökonomischer, eine Datenvalidierungsoftware gleich am Anfang der Kette zu implementieren.